The Reading Matrix: An International Online Journal
Volume 16, Number 2, September 2016

The Effects of Explicit Word Recognition Training on Japanese EFL Learners

Lance Burrows
Kindai University, Osaka, Japan

Michael Holsworth
Kyoto Sangyo University, Kyoto, Japan

ABSTRACT

This study is a quantitative, quasi-experimental investigation focusing on the effects of word recognition training on word recognition fluency, reading speed, and reading comprehension for 151 Japanese university students at a lower-intermediate reading proficiency level. Four treatment groups were given training in orthographic, phonological, and/or semantic processing, while a control group engaged in sustained silent reading. Treatment materials included two texts of different readability levels. Results indicated the intervention positively impacted participants' reading speed for both reading passages. However, reading comprehension and word recognition fluency improved more convincingly when learners were faced with reading passages that were slightly beyond their linguistic competence. The results highlight the importance of including a phonological element in word recognition training for students in foreign language reading classes.

INTRODUCTION

Research investigating reading comprehension skills has demonstrated that word recognition remains a powerful predictor of reading abilities not only in a L1 (e.g., Perfetti, Landi, & Oakhill, 2005) but also in L2 settings (e.g., Segalowitz & Hulstijn, 2005). However, there has been little research conducted to investigate specific ways in which readers can improve their word recognition skills. Attaining a high level of proficiency in word recognition remains paramount for learners attempting to become fluent L2 readers (Perfetti & Curtis, 1986), yet it is automatizing these skills that often poses them the most trouble. The present study focuses on a form of word recognition training and its corresponding effects.

The Verbal Efficiency Theory and Lexical Quality Hypothesis

The Verbal Efficiency Theory (VET) (Perfetti, 1985) assumes automatized and efficient word recognition skills as one of the core criteria for proficient reading comprehension. Evolving from the VET, the Lexical Quality Hypothesis (LQH) developed by Perfetti and Hart (2001) further states that word-recognition skills can become automatic when:

> …it (a word) has a fully specified orthographic representation (a spelling) and redundant phonological representations (one from spoken language and one recoverable from orthographic-to-phonological mappings). If a lexical representation is specific and redundant, its retrieval is more likely to be coherent and reliable. By coherent, we mean that the constituents are available synchronously at retrieval, giving the impression of a unitary word (Perfetti & Hart, 2001, p. 190).

The underlying premise of the LQH is that reliable, coherent, high-quality representations will be more efficiently retrieved than those of poor quality.

As alluded to earlier, there seems to be a strong connection between the efficiency of word recognition and text-level reading comprehension. According to the VET (Perfetti, 1985) and with strong support drawn from theories of automaticity of reading (LaBerge & Samuels, 1974), it can be extrapolated that text-level reading comprehension is enhanced as readers achieve improved levels of word recognition and reading speed. As readers become better able to quickly and accurately recognize words, they are also able to process strings of text faster, allowing more incoming information to remain in working memory longer (Perfetti, 1985), enabling the reader to make and maintain the necessary connections among ideas in the text that are integral to passage-level reading comprehension.

Empirical Research on Word Recognition Training

Inefficient phonemic decoding and word recognition skills have been shown to significantly impede accurate reading comprehension in L2 readers (Koda, 1996). Children learning to read in their L1 usually know several thousand words orally (Cunningham, 2005) so the main challenge for them is learning to decode, assigning phonological information to orthography. On the contrary, most students of L2 reading start with a significantly smaller repertoire of vocabulary, which means that not only must they become skillful at decoding words, but they must also learn the corresponding semantic information. Furthermore, these difficulties are compounded when the orthographies of the L1 and L2 are vastly distant, such as in the case of Japanese, a logographic/syllabary system, and English, an alphabetic system. L2 learners whose L1 is an alphabetic system differ from those whose L1 is a non-alphabetic system in orthographic processing (Ryan & Meara, 1991) as well as in phonological processing (Brown & Haynes, 1985; Koda, 1998).

In a study conducted in an EFL context, Akamatsu (2007) investigated the effects of word recognition training on word recognition processing. For the treatment, 49 first-year university students in Japan were given seven word-recognition training sessions over a seven-week period (i.e., one training session per week). In each session, which only lasted 90 seconds, participants were required to draw a line to separate strings of both high and low frequency words as quickly as possible. The results demonstrated that low frequency word training produced results with evidence of automatization of word recognition, whereas high frequency word training resulted in a speeding up of the word recognition process. Akamatsu claimed that the learners had already automatized the recognition of high frequency words so the improvements made were minimal and did not represent the qualitative changes that were attributed to the improvements in low frequency word recognition. Furthermore, the results from a reading comprehension test did not show any significant improvement in reading ability.

There are three possible reasons for the results on the reading comprehension measure. Firstly, the participants received only orthographic training on the target words with no explicit semantic or phonological training. Also, the target words used in the training process did not appear in the reading comprehension test, rendering a direct effect of word recognition training on text comprehension highly unlikely. Another possible explanation could have been the limited time-on-task for the treatment, only 10 minutes and 30 seconds in total.

In another related study, Fukkink, Hulstijn, and Simis (2005), investigated if the improvements in word recognition stemming from lexical access training would translate into improvements in reading speed and performance in a group of 105 L1 Dutch, grade 8 learners of English. The results showed that speed of lexical access improved; however, the effect on automatization was relatively weak. In addition, the apparent improvements in lexical access did not result in significantly faster reading speed or better reading performance even though all of the target words had been extracted from the passage used in the reading comprehension test. One possible explanation for this may be that during the reading comprehension test, the participants were not subject to a time constraint. This lack of time pressure could result in the use of reading strategies that potentially compensate for poor word recognition skills. In addition, the researchers did not include any explicit phonological training in the treatments.

These studies do not seem to address some of the crucial components of word recognition training and testing validity criteria. Firstly, the studies have offered no explicit training in phonological processing, despite the evidence that indicates it may be one of the most important components of word recognition ability (Torgesen & Burgess, 1998). Furthermore, there has been no intervention study that focuses on all three components of word recognition. In regard to reading comprehension, the testing methodology seems to suffer from possible flaws, including issues dealing with time-on-task. Lastly, none of the studies have considered how word recognition training might differentially impact learners when the target words are taken from passages of varying readability levels.

This study will attempt to address all of these issues by answering the following research questions:

RQ1. To what degree does engaging in word recognition treatments, or sustained silent reading for the control group, affect the participants' reading speed, word recognition fluency and reading comprehension? Hypothesis: Groups with more components of training, e.g., the OSP group, will outperform the others.

RQ2. Is there a differing effect from the word recognition training on the various groups depending on the level of the targeted reading passage (from which the target vocabulary were extracted)?

METHOD

Participants

The participants, all of whom were non-English, economics majors, consisted of 151 first-year Japanese university students enrolled in required reading classes. All of the participants were Japanese native speakers and as such, had studied English formally for at least six years in the Japanese public education system prior to the study. Five intact classes were randomly assigned to either one of the four experimental groups or the control group (see Table 1). The students in these five classes had a mean score of 278 on the reading section of the TOEIC™ test and a mean on a vocabulary size test of 3,460 words (Nation & Beglar, 2007).

Design

Experimental comparisons, conducted with a between-subjects design, involved the performance differences between five groups; a control group, a semantic group, an orthographic group, an orthographic/semantic group, and an orthographic/semantic/

Table 1. Review of Procedures for Each Group

Group	*n*	Tasks
control (C)	28	sustained silent reading in class
semantic (S)	26	generate original sentences using target words
orthographic (O)	37	matching and slash reading exercises
orthographic/ semantic (OS)	24	flashcards review (no chorusing), matching and slash reading exercises
orthographic/semantic /phonological (OSP)	36	chorusing the researcher, flashcard review (aloud), matching and slash reading exercises

phonological group (see Table 1). A reading test designed by Burrows (2012) was adapted and utilized to measure reading comprehension. Reading speed was calculated by recording the amount of time the participants took to read the passage on the reading test and word recognition fluency was measured through a lexical decision task. There were two cycles of word recognition training, each lasting four weeks.

Material Selection

The target words in this study were extracted from two texts that were part of a reading comprehension test created by Burrows (2012). The fifty target words taken from each reading were divided into 25 words from the first 1000 most frequent words in English (1K words) and 25 words from the second 1000 most frequent words, offlist words, and words from the Academic Word List (2K+ words). All the non-lexical proper nouns, e.g. Japan or United States, were recategorized as 1K words. The target words were divided into two groups because the results from the Akamatsu (2007) study had demonstrated that there was a difference in the way the participants improved in their word recognition of high frequency and low frequency words after word recognition training.

The text from which the target words for the first training cycle were taken contained 371 words, of which 278 were 1k words (tokens per type: 2.21) and 93 were 2K+ words (type-token ratio: 1.60). The Flesch reading ease score was 66.6 and the Flesch-Kincaid grade level was 9. The topic of the passage was sharks, one which most participants would have general knowledge about. The second text contained 373 words of which 318 were 1K words (type-token ratio: 2.24) and 55 words were 2K+ words (type-token ratio: 1.45). The Flesch reading ease score for the second passage was 51.6 and the Flesch-Kincaid grade level was 10. This passage was based on business management practices and because the students were economics majors, it was considered a topic that most of them would be familiar with.

Piloting revealed that the first text was at a challenging level for the students; the pilot study participants correctly answered an average of 81.3% (M =12.20, SD = 1.38) of the comprehension questions, with none of the participants achieving a perfect score. However, the second text was deemed exceedingly difficult for the students because the average comprehension score, 64.2% (*M* =9.31, *SD* = 1.86) was much lower than that of the first text. It was hypothesized that word recognition fluency training might start to show effect when learners were exposed to texts that were beyond their current reading proficiency level.

Treatment Procedures

The entire study lasted 8 weeks; two, four-week treatment cycles that focused on two different sets of target words. The time-on-task was controlled, 20 minutes per class for all five groups. The study was designed to maintain ecological validity, which meant that the participants were trained in their regular school setting during their normally-scheduled reading classes.

The *control* group engaged in sustained silent reading practice for twenty minutes during each class of the treatment periods. Extensive reading (ER) had been conducted in this class for several weeks prior to the study, so the learners had been trained in how to read extensively following Day and Bamford's ER guidelines (2002). In-class sustained silent reading was seen as a viable option for the control group because it would allow the students exposure to written text without the deliberate attention to any one of the components of word recognition. Students were instructed to read at least 3,000 words weekly from books in level one of the Oxford Bookworms graded reader series that were available at the university library. At the end of the eight-week training cycle, 98% of the students had read over the assigned 24,000 words, with the average words read being 25,453 words.

The *semantic* group focused on processing meaning only. The participants were required to generate original sentences based on the 50 target words from each treatment cycle. They were provided a bilingual word list and given 20 minutes to generate original sentences, all of which were submitted to the researchers. The students were told to start each week generating sentences from where they had left off on the list the previous week. All the students were able to create one sentence per word. If some students finished within the time limit, they were instructed to review and revise the sentences they had developed, instead of writing more sentences. Because the target words were presented to the participants in list form, the participants were arguably engaged in some orthographic processing. However, the group did not intentionally engage in repeated practice designed strictly to enhance the noticing of the orthography of the words.

The *orthographic* group practiced recognizing the orthography of the target words through matching and slash reading exercises. In the matching exercises (see Figure 1), the participants were shown a word either on paper or an overhead projector screen and were required to find and circle the correct answer from a set of possible answers on a worksheet (adapted from Grabe and Stoller, 2011). For the slash reading exercise (see Figure 1), the students were given 25 strings of words, each comprised of four target words written without any spaces between the words, and expected to draw a line where one word ended and the next one began (adapted from Akamatsu, 2007).

Figure 1. Examples of Matching Exercise and Slash Reading Exercise Items

Matching Exercise example item

1. **active**	action	achieve	activate	[active]

Slash Reading Exercise example item

1. dwindleconceivefiresaction	=	dwindle/conceive/fires/action

The *orthographic/semantic* group received treatment that was similar to that assigned to the orthographic group, but included an activity, flashcard review, designed to promote semantic processing practice in addition to the activities promoting orthographic processing. At the start of each class, the researcher distributed a bilingual list of all fifty target words that the participants could use to reference at any time during the treatment. Bilingual flashcards were then distributed to the students and they were instructed to look first at the English side of the flashcard and to recall the meaning in their L1. The use of flashcards for semantic processing practice was based on the training results found in Tan and Nicholson (1997). The students then engaged in the matching and slash reading exercises. These students never chorused the words on the list, nor did they practice articulating these words when doing the flashcard review.

The *orthographic/semantic/phonological* group received treatment in all three, word recognition components, including the phonological processing practice. At the beginning of the class, after the first author had distributed the bilingual word list, he chorused each word with the students three times. In addition, the students were required to say each target word aloud during the flashcard review, something the students had not done in any other group. During this phase of the treatment, students wore headphones and were required to talk into a microphone limiting interference from neighboring participants. After the chorusing and flashcard review, students engaged in the matching and slash reading exercises.

Instrumentation

At the start of each four-week treatment cycle, all the participants completed a word recognition pre-test, after which the 20-minute treatment period began. In weeks 2-4, the same 20-minute training schedule continued until finally in the fourth week, after the treatment session in that class, the students took the word recognition post-test and the reading comprehension test. All instruments are explained in detail in the following paragraphs.

Word recognition pre- and post-tests (pre-WRT/post-WRT)

These tests were performed on line using a lexical decision task format adapted from the website, www.lextutor.ca . The tests were administered at the beginning of each 4-week training cycle and again at the end. The stimuli were the 25, 1K words and the 25, 2K+ words, and 50 pronounceable non-words. All the target words were consistent in spelling-sound correspondences among English words.

The lexical decision task consisted of written instructions, 50 practice trials, and 100 test trials in that order. All practice trials and test trials were randomized automatically. For the analyses of these tests, both the reaction times and the rate of accuracy were recorded for both the 1K and 2K+ words. The data for the reaction times and accuracy rates was presented as gain scores by subtracting each student’s scores on the pre-test from those of the post-test.

Reading comprehension test/reading time

This test was adapted from Burrows (2012). On the original test, there were four questions based on vocabulary. This was reduced to three for each comprehension test. Also, three more inferencing questions were added to the text to offer a clearer understanding of the participant’s ability to infer meaning on a local and global text level. Each 4-week treatment cycle was concluded with a reading comprehension test conducted after the word-recognition post-test. Following each text were 15 questions, all of which were developed to measure the learners' ability to engage in higher level processing, such as inferencing local and global messages from the text (e.g., "What can be inferred about the author?"; "According to the passage, what can we infer about sharks?"). The texts from these tests were the texts from which the target vocabulary were extracted. Cronbach's α for the first and second reading comprehension tests were .97 and .96, respectively.

Based on piloting, time limits for the first reading test was set at 11 minutes and 30 seconds and the second at 12 minutes and 15 seconds. When answering the comprehension questions, students were allowed to look back at the text. In addition to the questions, each student recorded his own reading speed by referring to an online stopwatch that was projected on an overhead screen.

DATA ANALYSIS AND RESULTS

The objective of this study was to investigate the effects of word recognition training on word recognition fluency, reading speed, and reading comprehension. Because the treatment cycles only lasted four weeks, it was not considered reliable to offer a pre- and post-reading comprehension test due to a possible test effect. Therefore, two covariates were used to adjust for any differences in reading proficiency or vocabulary knowledge from the outset of the study. The scores from a vocabulary size test were used as the covariate for vocabulary knowledge, while the scores from the TOEIC™ reading section were used as the covariate for reading proficiency.

To answer the research questions, data for the first and second reading passages were analyzed separately. One-way MANCOVA analyses were conducted to analyze the data for the variables: reading times, the reading comprehension scores, and the word recognition times and rates of accuracy for both 1K and 2K+ words. The independent variable was the five groups. (see Table 2 for descriptive statistics of passage 1, Table 4 for passage 2).

Reading Passage 1

A preliminary analysis evaluating the homogeneity-of-slopes assumption indicated that the relationship between the covariates and each of the dependent variables did not differ significantly ($p < 0.05$) as a function of the independent variable. In addition, the partial η^2 for each of the dependent variables was extremely low (all > 0.09) for both covariates.

Further preliminary analyses showed that Levene's test of equality of error variance was insignificant for all variables but the ReacTime1K variable. Because neither Brown-Forsythe F nor Welch's F statistics that normally would be used to further investigate the homogeneity of between-group variance in an analysis of variance (ANOVA), is available for MANCOVA analyses, it was decided to continue with the analyses and deal with the violation of homogeneity on the ANCOVA level with the problematic dependent variable. The results of the MANCOVA showed a significant multivariate effect for the combined dependent variables in respect of the group factor. Because there were unequal group sizes and the assumption-of-homogeneity of covariance matrices was not violated, Pillai's trace (V) was used (Field, 2005). The Pillai's V of .66 is significant, $F(24, 484) = 3.96$, $p < .01$. The multivariate η^2 based on Pillai's V was moderate, .17. The descriptive statistics for each of the dependent variables for each group are provided in Table 2.

In order to determine more clearly where the differences occurred, analyses of covariance (ANCOVA) on each dependent variable were conducted as follow-up tests to the MANCOVA (see Table 3). Using the Bonferroni method, each ANCOVA was tested at the $p < .005$ significance level, because there were 10 possible comparisons (Field, 2005). Due to the violation in equality of error variance, the results for the ANCOVA for the ReacTime1K variable must be interpreted cautiously.

Table 2. Descriptive Statistics for Reading Passage 1

	Reading Comprehension					Reading Time				
	Control	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho	Control	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho
M	-.51	-.52	.02	.04	-.01	205.45	239.26	169.16	192.17	187.39
SE	.17	.21	.17	.19	.16	12.22	17.49	8.72	9.61	8.93
95% CI										
LB	-.86	-.97	-.31	-.38	-.33	179.88	202.52	151.48	176.46	169.26
UB	-.16	-.07	.36	.46	.31	231.02	276.01	186.85	208.88	205.51
SD	.75	.93	1.00	.84	.95	54.65	76.23	53.04	55.73	53.57

	Gains in Accuracy of 1K Words					Gains in Accuracy of 2K+ Words				
	Control	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho	Control	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho
M	.15	-.16	.11	.00	-.03	.10	2.58	5.44	5.54	6.17
SE	.11	.14	.09	.08	.10	.45	.57	.59	.56	.59
95% CI										
LB	-.08	-.45	-.06	-.17	-.23	-.85	1.37	4.20	4.41	4.98
UB	.38	.13	.28	.17	.18	1.05	3.78	6.69	6.67	7.36
SD	.49	.60	.52	.34	.61	2.02	2.50	2.50	3.38	3.52

	Gains in Reaction Time of 1K Words					Gains in Reaction Time of 2K+ Words				
	Control	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho	Control	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho
M	.17	-.12	.30	.34	.35	.33	.45	.71	.83	.87
SE	.05	.09	.04	.06	.10	.08	.08	.07	.08	.09
95% CI										
LB	.07	-.07	.22	.21	.14	.16	.28	.57	.67	.70
UB	.27	.30	.380	.47	.56	.50	.62	.86	.99	1.05
SD	.22	.39	.25	.26	.63	.37	.35	.44	.48	.36

Note. Ortho/Sem = orthographic/semantic group; Or/Sem/Pho = orthographic/semantic/phonological group

Table 3. ANCOVA Results for Reading Passage 1

Dependent variable	df	Error	F	*MSE*	*p*	η^2
ReadComp	4	123	2.35	.79	.062	.07
ReadTime*	4	123	5.35	3175.97	.000	.15
Acc1K	4	123	1.13	.29	.333	.04
ReacTime1K[a]	4	123	1.35	.17	.234	.05
Acc2K+*	4	123	18.01	8.70	.000	.37
ReacTime2K+*	4	123	6.81	.18	.001	.18

Note. ReadComp = reading comprehension; ReadTime = reading time; Acc1K = gains in accuracy for 1K; ReacTime1K = gains in reaction time for 1K; Acc2K+ = gains in accuracy for 2K+; ReacTime2K+ = gains in reaction time for 2K+; ([a]) = did not meet assumption of equality in error variance; (*) $p < .005$.

Follow-up, pair-wise contrasts were conducted to evaluate differences among the adjusted means for those variables with significant ANCOVAs: 1) ReadTime, 2) Acc2K+, and 3) ReacTime2K+. Again, the Bonferroni adjustment was made; significance set at $p < .005$.

For the ReadTime variable (see Table 2), the significant differences existed between the control (C) and orthographic (O) groups, the semantic (S) and O groups, and the S and orthographic/semantic/phonological (OSP) groups. Therefore, the O and the OSP groups benefited most from the word recognition treatments in regard to reading speed.

In regard to the gains in Acc2K+ variable (see Table 2), significant differences could be found between the C and all three O, orthographic/semantic (OS), and OSP groups; and the S and all three O, OS, and OSP groups. Therefore, the groups that focused on the word recognition treatments with the explicit orthographic element outperformed the control and semantic groups.

For the gains in ReacTime2K+ variable (see Table 2), significant differences existed between the C and all three O, OS, and OSP groups; and the S and OS and OSP groups. Similar to the earlier results, the groups that focused on the word recognition treatments with the orthographic element convincingly outperformed the control and semantic groups.

Overall, the statistics indicate that although the word recognition treatment did not allow the experimental groups to significantly outperform their control group counterparts on the reading comprehension test, some variables, i.e., reading time, accuracy and reaction time of the 2K+ words did improve due to the treatment for the first reading passage.

Reading Passage 2

A preliminary analysis evaluating the homogeneity-of-slopes assumption indicated that the relationship between the covariates and two of the dependent variables differed significantly ($p < 0.05$) as a function of the independent variable. For the TOEIC™ reading section covariate, the ReadTime variable violated the assumption. For the VST covariate, the gains in Acc1K variable violated the assumption. Because there were only two dependent variables that presented problems for the covariates, the MANCOVA analyses were conducted with the intention of dealing with the problematic covariates in the follow-up ANCOVA stage. The descriptive statistics for each of the dependent variables for each group are provided in Table 4.

Further preliminary analysis of the MANCOVA results showed that Levene's test of equality of error variance was insignificant for all variables. The results of the MANCOVA showed a significant multivariate effect for the combined dependent variables in respect of the group factor. The Pillai's *V* of .81 is significant, $F(24, 492) = 5.04$, $p < .01$. The

Table 4. Descriptive Statistics for Reading Passage 2

	Reading Comprehension					Reading Time				
	Control	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho	Control	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho
M	-.41	-.40	.00	.45	.45	233.15	210.22	182.11	199.06	177.50
SE	.21	.13	.11	.16	.10	11.17	15.15	9.52	10.78	7.36
95% CI										
LB	-.85	-.66	-.22	.13	.25	209.76	178.39	162.80	176.31	162.56
UB	.03	-.13	.23	.78	.64	256.54	242.57	201.42	221.80	192.44
SD	.94	.55	.67	.66	.58	49.97	66.03	57.91	45.73	44.16

	Gains in Accuracy of 1K Words					Gains in Accuracy of 2K+ Words				
	Control	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho	Control	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho
M	.12	.02	.08	-.13	-.01	-.10	1.63	3.16	2.94	3.06
SE	.10	.09	.09	.14	.10	.27	.22	.42	.41	.30
95% CI										
LB	-.04	-.14	-.08	-.38	-.21	-.66	1.16	2.31	2.08	2.44
UB	.28	.19	.24	.11	.19	.47	2.09	4.01	3.81	3.67
SD	.34	.28	.46	.53	.47	1.21	.96	2.54	1.73	1.82

	Gains in Reaction Time of 1K Words					Gains in Reaction Time of 2K+ Words				
	Control	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho	Control	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho
M	.12	.33	.22	.21	.15	.07	.19	.40	.33	.29
SE	.04	.08	.03	.05	.04	.09	.08	.04	.05	.03
95% CI										
LB	.04	.16	.15	.11	.07	-.11	.02	.32	.21	.22
UB	.20	.50	.28	.31	.23	.26	.37	.48	.44	.36
SD	.17	.36	.20	.20	.23	.39	.36	.24	.23	.20

Note. Ortho/Sem = orthographic/semantic group; Or/Sem/Pho = orthographic/semantic/phonological group.

multivariate η^2 based on Pillai's V was moderately strong, .20. The descriptive statistics for each of the dependent variables for each group are provided in Table 5.

As a follow-up to the significant MANCOVA, respective ANCOVAs were conducted on each dependent variable; significance set at $p < .005$. Because the TOEIC™ covariate violated the homogeneity-of-slopes assumption for the ReadTime variable, it was removed from the ANCOVA for that variable. Similarly, because the VST covariate violated the assumption for the Acc1K variable, it was removed from the ANCOVA for that variable.

Because all but two of the ANCOVAs (see Table 5), Acc1K and ReacTime1K variables, resulted in statistical significance, pair-wise contrasts were conducted to evaluate differences among the adjusted means for the other four variables, $p < .05$. For the ReadComp variable, the OS and OSP groups outperformed all three other groups, the C, S, and O groups. For the ReadTime variable, the OSP group was significantly faster than the C and S groups, and the O group was faster than the C group. For the gains in Acc2K+ variable, the O, OS, and OSP groups outperformed the C group, and the OSP group outperformed the S group. Finally, in regard to the gains in ReacTime2K+ variable, the O, OS, and OSP groups all outperformed the C group. (For a summary of the follow-up ANCOVA results of both reading passages 1 and 2, see Table 6.)

Table 5. ANCOVA Results for Reading Passage 2

Dependent variable	df	Error	F	*MSE*	*p*	η^2
ReadComp*	4	123	8.51	.45	.000	.22
ReadTime[a]*	4	124	4.42	2776.71	.002	.13
Acc1K	4	123	1.17	.22	.381	.06
ReacTime1K	4	123	2.45	5.46	.050	.07
Acc2K+*	4	125	12.56	3.55	.000	.29
ReacTime2K+[b]*	4	124	4.91	7.65	.001	.14

Note. ReadComp = reading comprehension variable; Acc1K = gains in accuracy for 1K; ReacTime1K = gains in reaction time for 1K; Acc2K+ = gains in accuracy for 2K+; ReacTime2K+ = gains in reaction time for 2K+; ([a]) = TOEIC™ covariate omitted from analysis; ([b]) = VST covariate omitted from analysis; (*) $p < .005$.

Table 6. Summary of Variables with Significant Differences for Reading Passages 1 and 2

	Reading Passage 1			
	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho
Control		ReadTime Acc2K+ ReacTime2K+	Acc2K+ ReacTime2K+	Acc2K+ ReacTime2K+
Semanti c		ReadTime Acc2K+	Acc2K+	ReadTime Acc2K+

	Reading Passage 2			
	Semantic	Orthographic	Ortho/Sem	Or/Sem/Pho
Control		ReadTime Acc2K+ ReacTime2K+	ReadComp Acc2K+ ReacTime2K+	ReadComp ReadTime Acc2K+ ReacTime2K+
Semanti c			ReadComp	ReadComp ReadTime Acc2K+
Ortho			ReadComp	ReadComp

Note. The outperforming group names run horizontally along the top of the table. The underperforming group names run vertically along the left edge of the table. Ortho = orthographic group; Ortho/Sem = orthographic/semantic group; Or/Sem/Pho = orthographic/semantic/phonological group; effect size values are offered in Tables 3 and 5.

DISCUSSION

Research Question 1

In order to answer this question, results for the two reading passages will be addressed separately at first, and the relevant information will be synthesized at the end for a final analysis.

Reading passage 1

The hypothesis for this question stated the groups that received training on more components of word recognition (i.e., orthographic, semantic, and phonological information) would outperform those groups that focused on fewer components. This hypothesis is only partially supported by the results. Firstly, the results for the ReadComp, Acc1K, ReacTime1K variables were not significantly different between the groups. In regard to the Acc1K and ReacTime1K variables, this statistical insignificance was also exhibited in the Akamatsu study (2007).

In stark contrast, the word recognition fluency (both Acc2K+ and ReacTime2K+) for 2K+ words showed significant gains for the groups that focused on only orthographic processing or a combination of orthographic processing and the other elements (i.e., O, OS, OSP groups). These results also mirror those of Akamatsu (2007) in which the word recognition of low frequency words was thought to have undergone a qualitative change due to the intervention in his study.

The O, OS, and OSP groups participated in rigorous practice activities that incorporated one or more of the three word recognition elements. Presumably, word recognition skills follow the same kind of power function that has been identified in the acquisition of foreign language listening and speaking skills (DeKeyser, 1997) and lexical recognition processes in reading (Kirsner, 1994). Therefore, deliberate attention to these aspects of word recognition and sufficient, consistent practice may be necessary, especially in regard to 2K+ vocabulary.

Reading passage 2

Regarding the first research question, the results for the second reading passage more clearly support the hypothesis that those groups who received training in more components of word recognition would outperform those groups that had received training on fewer components. As shown in Table 6, the Acc1K, ReacTime1K, Acc2K+ and ReacTime1K variables seemed to behave in nearly exactly the same way for reading passage 2 as they did for reading passage 1, but the ReadComp and ReadTime variables exhibited different properties in reading passage 2 and deserve further explanation.

The first obvious justification for the significant differences in reading comprehension between the OS and OSP groups, and the C, S and O groups would be the interventions that the various groups had received. Along with the practice activities on orthography, the OS and OSP groups focused on the lexical meaning of the target words by using flashcards in class. The OSP group was additionally given explicit instruction on pronunciation of the words through chorusing with the instructor and pronouncing the target words while performing the flashcard review. The results regarding the use of flashcards were consistent with the findings of Tan and Nicholson (1997) in which the learners improved in reading comprehension, reading speed and accuracy after orally reciting target words.

Because the participants in this study were Japanese, the phonological element of the OSP intervention was particularly relevant. As explained in the literature review, Japanese learners face distinct obstacles in phonological processing (e.g., Brown & Haynes, 1985). This phonological component was lacking in previous studies (Akamatsu, 2007; Fukkink et al., 2005), which may have contributed to the insignificant reading comprehension results of those studies. The cocktail of activities assigned to the OSP group resulted in the participants acquiring a coherent representation of the target words. According to the Lexical Quality Hypothesis (Perfetti & Hart, 2001), these types of high-quality representations will be more efficiently retrieved and possibly assist in reading comprehension.

The flashcard review utilized in the OS and OSP groups seemed to produce superior results over those of the S group that had been required to generate sentences from the target words. One possible explanation for this is that during the flashcard review, participants in the OS and OSP groups were not only repeatedly exposed to the semantic information of the target words but the orthographic information as well. That is, they were required to practice the identification as well as the lexical access of the words. As mentioned earlier, the S group participants were allowed to see the orthography of the words but were not required to attend to it. They were only asked to focus on the meaning of the words. According to several researchers, an explicit focus on form, in this case orthography, can be useful in the learning process (DeKeyser, 1995). Conversely, the participants of the O group were primarily focusing on the orthography of the words and not given an explicit opportunity to practice accessing the meanings. This could help explain why the participants of this group were able to outperform the C group on accuracy and reaction time of 2K+ words with no appreciable gains in reading comprehension. They were able to orthographically identify words, but may have had trouble with the lexical access of certain words in the text.

Research Question 2

Research question 2 examined whether there was a differing effect from the word recognition training on the various groups depending on the level of the reading passage (from which the target vocabulary were extracted). In order to answer this question the ANCOVA results for both reading passages must be re-inspected (see Table 6). Preliminary analysis shows that there were 12 significant differences among the five groups for reading passage 1 and 16 differences for reading passage 2. In comparison to the C group, for both reading passages 1 and 2, the Acc1K and ReacTime1K, and the Acc2K+ and ReacTime2K+ variables behaved similarly so there does not seem to be very much difference for these variables depending on reading passage difficulty.

However, obvious differences start to emerge with further inspection of the ReadTime and ReadComp variables. For reading passage 1, the O group was the fastest group and achieved significant differences in comparison to the C and S groups. For reading passage 2, the OSP group attained the best reading time. A possible argument to explain this is that the first reading passage was at a level consistent with the participants' reading proficiency level and therefore, the intervention was fairly ineffective in generating significant differences among the groups. As the level of the second reading passage was slightly above the learners' proficiency level, the intervention proved more useful to the participants in the experimental groups.

A similar, but more pronounced trend was evident with the ReadComp variable. For the first passage, none of the treatment groups were able to outperform their control group counterparts, however, in the second cycle of treatment, there were six significant comparisons. As the difficulty level of the reading passage increased, the groups that seemed to benefit from the treatments were the OS and OSP groups, outperforming all other three groups. Because the readability of the second passage was lower than that of the first passage, it can be argued that the participants in the OS and OSP groups benefitted from the extensive training on the phonological, orthographic, and semantic aspects of the target words.

CONCLUSION

Pedagogical Implications

The results of this study highlight the potential that word recognition training holds for the foreign language classroom. One of the recurring themes was the importance of explicit attention being paid to all three of the word recognition elements. Another issue that appeared was the call for consistent practice in regard to these types of activities. Word recognition training should not be intermittently introduced but rather presented on a regular basis in tandem with any focus on vocabulary that might be implemented in a reading program. Also, in order to keep students motivated and interested in the activities, practitioners should consider varying the exercises that are offered to learners. Challenging the learners with a time constraint will also keep the learners engaged and stimulated.

Future Research

This study raises a host of questions concerning word recognition:

- In this study, fifty words were highlighted in each four-week cycle. Is there a certain number of words and/or time frame that should be included for optimal achievement?
- This study introduced an explicit approach to training word recognition. Are there other methods, possibly that focus on implicit training of word recognition that might assist learners in honing their word recognition skills, (e.g. extensive reading)?
- What might the long-term effects of consistent word recognition practice be on word recognition skills and reading comprehension?

Limitations

One issue that weakened the study was that there was no instrument, other than the reading comprehension test, to test the participants' semantic or phonological knowledge of the target words. In the word recognition tests, only the orthographic identification of the target words was measured. In order to gain a more accurate understanding of the effects of word recognition training, instruments to measure all three aspects of word recognition should have been included.

One other possible criticism of the study might be the way in which the experimental groups were organized. In order to determine the value of explicit training on each or combination of the three components of word recognition, groups were formed to focus on certain elements. The explicit training of each of the components can in no way truly isolate the cognitive processes underlying that component. For example, when words were presented to the orthographic group, although the participants were being explicitly introduced and expected to practice the orthographic identification of the words, it would be irresponsible to claim that other processes were not being initiated automatically due to the participant's interaction with the orthography. Phonological and possibly semantic processing would also commence. The objective of the study was not to isolate underlying cognitive processes but to determine to what degree, explicit training focusing on strengthening the various components of word recognition might affect overall word recognition ability and reading comprehension.

Lance Burrows *is an Associate Professor at Kindai University in Osaka, Japan where he teaches mainly reading and writing courses. His research interests include the effects of extensive reading and reading strategies on affective factors such as reading self-efficacy and intrinsic motivation.*

Email: lanceburrows@hotmail.com

Michael Holsworth *is an assistant professor at Kyoto Sangyo University in Kyoto, Japan. His research focuses largely on exploring ways to help learners improve their word recognition skills and overall reading comprehension. He is also interested in curriculum design using a learner-centered teaching approach.*

Email: mikesensei007@gmail.com

REFERENCES

Akamatsu, N. (2007). The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics*, *29*, 175–193.

Brown, T., & Haynes, M. (1985). Literacy background and reading development in a second language. In H. Carr (Ed.), *The development of reading skills* (19–34). San Francisco, CA: Jossey-Bass.

Burrows, L. (2012). *The effects of extensive reading and reading strategies on reading self-efficacy*. Unpublished dissertation. Temple University, Osaka, Japan.

Cunningham, A. (2005). Vocabulary growth through independent reading and reading aloud to children. In E. Hiebert & M. Kamil (Eds.), *Teaching and learning vocabulary* (45–68). Mahwah, NJ: Lawrence Erlbaum.

Day, R. R., and Bamford, J. (2002). Top ten principles for teaching extensive reading. *Reading in a Foreign Language, 14*, 136–141.

DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition, 17*, 249–297.

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19*, 19–221.

Fukkink, R. G., Hulstijn, J., & Simis, A. (2005). Does training in second□language word recognition skills affect reading comprehension? An experimental study. *The Modern Language Journal*, *89*, 54–75.

Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading* (2nd ed.). Harlow, UK: Pearson Education.

Kirsner, K. (1994). Implicit processes in second language learning. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (283–311). London: Academic Press.

Koda, K. (1996). L2 word recognition research: A critical review. *The Modern Language Journal*, *80*, 450–460.

Koda, K. (1998). The role of phonemic awareness in second language reading. *Second Language Research, 14*, 194–215.

LaBerge, D., & Samuels, J. (1974). Towards a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293–323.

Linacre, J. M. (2006). WINSTEPS: Rasch model computer program [Computer software]. Chicago: Winsteps.com.

Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, *31*, 9–12.

Perfetti, C. A. (1985). *Reading ability.* New York: Oxford University Press.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*. *11*, 357–383.

Perfetti, C. A., & Hart, L. (2001). The lexical bases of comprehension skill. In D. Gorfien (Ed.), *On the consequences of meaning selection* (67–86). Washington, DC: American Psychological Association.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling, & C. Hulme (Eds.), *The science of reading: A handbook* (227–247). Oxford, UK: Blackwell Publishing.

Perfetti, C. A. & Curtis, M. (1986). Reading. In R. F. Dillon & R. J. Sternberg (Eds.), *Cognition and instruction* (13–57). New York: Academic Press.

Ryan, A., & Meara, P. (1991). The case of the invisible vowels: Arabic speakers reading English words. *Reading in a Foreign Language, 7*, 531–540.

Segalowitz, N. & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (371–388). Oxford, UK: Oxford University Press.

Tan, A., & Nicholson, T. (1997). Flashcards revisited: Training poor readers to read words faster improves their comprehension of text. *Journal of Educational Psychology*. *89*, 276–288.

Torgesen, J. K., & Burgess, S. R. (1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal-correlational and instructional studies. In J. Metsala & L. Ehri (Eds.). *Word recognition in beginning reading*. Hillsdale, NJ: Lawrence Erlbaum.

Wright, B. D., Linacre, J. M., Gustafson, J. E., & Martin-Lof, P. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8,* 370. *Last accessed July 21, 2015 from* http://rasch.org/rmt/rmt83b.htm